

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 19:39:28 ; Search time 115 seconds
(without alignments)
1216.560 Million cell updates/sec

Title: US-09-717-789B-2

Perfect score: 2122

Sequence: 1 MALVNLVHGHTSEKQIQI.....PWEKENLSDGDFEDANKEQ 390

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2122	100.0	390	3 AAY58158	AAY58158 Adeno ass
2	2122	100.0	390	5 AAU11403	AAU11403 Adeno-ass
3	2122	100.0	610	3 AAY58159	AAY58159 Adeno ass
4	2122	100.0	610	4 AAY97720	AAY97720 Rep prote
5	2122	100.0	610	5 AAU11404	AAU11404 Adeno-ass
6	2122	100.0	610	5 AAE22887	Aae22887 Adeno-ass
7	2122	100.0	610	5 AAE22864	Aae22864 Adeno-ass
8	2122	100.0	610	5 AAE22694	Aae22694 Adeno-ass
9	2122	100.0	610	6 ABU64865	ABU64865 Rep prote
10	2122	100.0	610	6 ABU64760	ABU64760 Adeno ass
11	2122	100.0	610	6 ABR43398	ABR43398 Adeno-ass
12	2122	100.0	610	7 ADI40280	ADI40280 Adeno-ass
13	2122	100.0	610	8 ADH58895	ADH58895 Predeterm
14	1721	81.1	330	3 AAY58163	AAU11408 Adeno-ass
15	1721	81.1	330	5 AAU11408	AAU11408 Adeno-ass
16	1721	81.1	550	3 AAY58164	AAY58164 Adeno ass
17	1721	81.1	550	5 AAU11409	AAU11409 Adeno-ass
18	1232.5	58.1	399	2 AAU46310	AAU46310 AAU4 Rep
19	1232.5	58.1	399	2 ABR73937	ABG73937 Adeno-ass
20	1232.5	58.1	623	6 AAW46307	AAW46307 AAU4 Rep
21	1232.5	58.1	623	2 AAU46312	AAU46312 AAU4 Rep
22	1232.5	58.1	623	2 AAY97712	AAY97712 Rep prote
23	1232.5	58.1	623	5 AAE28636	Aae28636 Adeno-ass
24	1232.5	58.1	623	5 AAE26932	Aae26932 Adeno ass
25	1232.5	58.1	623	6 ABU64857	ABU64857 Rep prote

26	1232.5	58.1	623	6 ABU64752	ABU64752 Adeno ass
27	1232.5	58.1	623	6 ABR73939	ABG73939 Adeno-ass
28	1232.5	58.1	623	6 ABR73934	ABG73934 Adeno-ass
29	1232.5	58.1	623	6 ABR43390	ABR43390 Adeno-ass
30	1232.5	58.1	623	8 ADH58893	ADH58893 Predeterm
31	1231	58.0	623	6 ABB80230	ABR80230 AAV9 rep
32	1229	57.9	623	7 ABR62760	ABR62760 Adeno ass
33	1227.5	57.8	623	7 ADI40264	ADI40264 Adeno-ass
34	1226	57.8	623	7 ADE76504	ADe76504 Adeno-ass
35	1223.5	57.7	623	4 AAY97713	AAY97713 Rep78 pro
36	1223.5	57.7	623	5 AAE22880	Aae22880 Adeno-ass
37	1223.5	57.7	623	5 AAE28637	Aae28637 Adeno-ass
38	1223.5	57.7	623	5 AAE26933	Aae26933 Adeno ass
39	1223.5	57.7	623	6 ABU64858	ABU64858 Rep78 pro
40	1223.5	57.7	623	6 ABR43391	ABR43391 Adeno ass
41	1223.5	57.7	623	7 ADI40266	ADI40266 Adeno-ass
42	1223.5	57.7	623	7 ADI40266	ADI40266 Adeno-ass
43	1223.5	57.7	624	4 AAB59850	AAB59850 AAV3B Rep
44	1223.5	57.7	624	8 ADH58891	ADH58891 Predeterm
45	1223.5	57.7	624	8 ADH58892	ADH58892 Predeterm

ALIGNMENTS

RESULT 1
AAY58158

ID AAY58158 standard; protein; 390 AA.

AC AAY58158;

DT 07-MAR-2000 (first entry)

DE Adeno associated virus AAV5 Rep52 protein.

XX Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR;

KW Promoter; Rep protein; capsid protein; regulation; transcription;

KW Replication; chromosomal integration; tissue tropism; cellular receptor;

KW gene therapy; neutralising antibody; erythroid progenitor cell;

XX transduction; cancer; genetic disease; Rep52.

OS Adeno-associated virus 5.

XX WO9961601-A2.

PN 02-DEC-1999.

PF 28-MAY-1999; 99WO-US011958.

PR 28-MAY-1998; 98US-0087029P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Chlorini JA, Kotin RM;

XX WPI; 2000-062707/05.

DR N-ESDB; AA49210.

XX Adeno-associated virus 5 based vectors and particles, useful for gene therapy.

XX Claim 14; Page 75; 91pp; English.

This sequence represents the Rep52 protein of adeno associated virus type 5 (AAV5). The invention relates to vectors comprising a pair of AAV5 inverted terminal repeats (ITRs) with a promoter between the ITRs. The vector may comprise the viral genome, or subregions thereof, including sequences encoding Rep proteins and capsid proteins, and is encapsidated in an AAV5 particle. The non-structural Rep proteins Rep40 (AAY58613), Rep52 (AAY58168), Rep68 (AAY58164) and Rep78 (AAY58159) are involved in regulation of replication and transcription, in addition to the production of progeny genomes. Rep68 and Rep78 are also associated with the stable integration of the viral genome into human chromosomes. The

Db 401 CKSSVQIDSTPVIIVTNTNMCVVVDGNTTFFHQOPLDRMFKELTKRLPPDFGKITKQ 460
 QY 241 EVKDFFAWAKVQVPTHEFKVPRELAGTGAESLKRPLGDVNTNTSYKSLEKARLSFV 300
 Db 461 EVKDFFAWAKVQVPTHEFKVPRELAGTGAESLKRPLGDVNTNTSYKSLEKARLSFV 520
 QY 301 PETPRSSDVTVDPAPLRPLNWSRYDCKDCHAOFDNISNKCDECVLNKRGKNCICHNV 360
 Db 521 PETPRSSDVTVDPAPLRPLNWSRYDCKDCHAOFDNISNKCDECVLNKRGKNCICHNV 580
 QY 361 THCQICHGIPPWEKENLSDFGDFDANKEQ 390
 Db 581 THCQICHGIPPWEKENLSDFGDFDANKEQ 610

RESULT 6

AAE22887
 ID AAE22887 standard; protein; 610 AA.

XX AAE22887;

AC AAE22887;

DT 29-AUG-2003 (revised)

DT 09-AUG-2002 (first entry)

XX

DE Adeno-associated virus 5 Rep protein.

XX Nucleic acid/protein conjugate; NAP; nucleic acid modification; NAM; EAS;
 KW enzyme attachment sequence; cancer therapy; protein-protein interaction;
 KW drug discovery; Rep protein; adeno-associated virus; AAV; gene therapy;
 KW cytostatic; Rep protein.

XX

OS Adeno associated virus; 5.

XX

XX WO200222826-A2.

XX

PD 21-MAR-2002.

XX

PF 14-SEP-2001; 2001WO-US028702.

XX

PR 14-SEP-2000; 2000US-0232960P.

XX

XX (XENC-) XENCOR INC.

XX

XX Li M, Melander C, Liu H;

XX

XX WPI; 2002-393969/42.

XX

XX N-PSDB; AAD36281.

XX

PT Library of nucleic acid/protein conjugates, has a fusion of nucleic acid
 PT modification enzyme and candidate compound, and expression vector having
 PT a fusion of nucleic acids encoding NAM enzyme and the compound.

XX

PS Disclosure; Fig 21; 96pp; English.

XX

CC The present invention relates to genetic libraries of nucleic acid/
 CC protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic
 CC acid modification (NAM) enzyme (E) and candidate compound), an expression
 CC vector (with a fusion of nucleic acids encoding the enzyme and candidate
 CC protein respectively), an enzyme attachment sequence (EAS; RNA sequence),
 CC where the candidate compound and candidate protein are different and EAS
 CC and the enzyme are covalently linked. The NAP conjugates are useful in
 CC screens to assay binding to target molecules and/or to screen candidate
 CC agents for the ability to modulate the activity of the target molecule.
 CC They are useful in cancer therapy. Sequences of the invention are also
 CC useful to detect protein-protein interaction, in drug discovery, to
 CC discover DNA or nucleic acid binding proteins, using nucleic acids as the
 CC targets and to screen for NAM enzymes with decreased toxicity for host
 CC cells (specifically Rep proteins with reduced toxicity). NAP conjugates
 CC are also useful in pharmacogenomic studies, for screening bioactive
 CC agents on surface cells, viruses and microbial organisms. They are also
 CC useful for screening proteins causing phenotypic changes such as
 CC overproduction or inhibition of protein expression, or proteins that

CC alter attachment, infectivity, etc. of the virus. Sequences of the
 CC invention are also used in gene therapy. The present sequence is adeno-
 CC associated virus (AAV) 5 Rep. (Updated on 29-AUG-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 610 AA;

Query Match 100.0%; Score 2122; DB 5; Length 610;
 Best Local Similarity 100.0%; Pred. No. 8 6e-188;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVNMVLVEHGITSEKQWIOENQESYLSFNSTGNSRQIKAAALDNATKIMSLTKSAVDYL 60
 Db 221 MALVNMVLVEHGITSEKQWIOENQESYLSFNSTGNSRQIKAAALDNATKIMSLTKSAVDYL 280
 QY 61 VGSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKNKTWLYGPATTGKTNIA 120
 Db 281 VGSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKNKTWLYGPATTGKTNIA 340
 QY 121 EAJAHTVPFYGCYVNTNENPFNDKMLIWMEEGKMTNKVVESAKAILGGSKVRVDOK 180
 Db 341 EAJAHTVPFYGCYVNTNENPFNDKMLIWMEEGKMTNKVVESAKAILGGSKVRVDOK 400
 QY 181 CKSSVQIDSTPVIIVTNTNMCVVVDGNTTFFHQOPLDRMFKELTKRLPPDFGKITKQ 240
 Db 401 CKSSVQIDSTPVIIVTNTNMCVVVDGNTTFFHQOPLDRMFKELTKRLPPDFGKITKQ 460
 QY 241 EVKDFFAWAKVQVPTHEFKVPRELAGTGAESLKRPLGDVNTNTSYKSLEKARLSFV 300
 Db 461 EVKDFFAWAKVQVPTHEFKVPRELAGTGAESLKRPLGDVNTNTSYKSLEKARLSFV 520
 QY 301 PETPRSSDVTVDPAPLRPLNWSRYDCKDCHAOFDNISNKCDECVLNKRGKNCICHNV 360
 Db 521 PETPRSSDVTVDPAPLRPLNWSRYDCKDCHAOFDNISNKCDECVLNKRGKNCICHNV 580
 QY 361 THCQICHGIPPWEKENLSDFGDFDANKEQ 390
 Db 581 THCQICHGIPPWEKENLSDFGDFDANKEQ 610

RESULT 7

AAE28644
 ID AAE28644 standard; protein; 610 AA.

XX AAE28644;

AC AAE28644;

DT 29-AUG-2003 (revised)

DT 27-DEC-2002 (first entry)

XX

DE Adeno-associated virus 5 Rep protein.

XX

XX Nucleic acid modification enzyme; NAM; enzyme attachment sequence; EAS;

XX protein design automation; PDA; cancer; protein-protein interaction;

XX infection; Gene therapy; Rep protein.

XX

OS Adeno associated virus; 5.

XX

XX WO200268453-A2.

XX

XX 06-SEP-2002.

XX

XX 19-FEB-2002; 2002WO-US004853.

XX

XX 22-FEB-2001; 2001US-00792629.

XX

XX (XENC-) XENCOR INC.

XX

XX Li M, Dahiyat BI;

XX

XX WPI; 2002-691653/74.

XX

XX N-PSDB; AAD46138.

XX

PT Generating a library of fusion nucleic acids for treating cancer or

Mon Oct 25 07:55:19 2004

us-09-717-789b-2.rag

PT infection, or detecting protein-protein interaction, comprises providing
PT computationally-derived library of candidate protein sequences and
PT expression vectors.

XX Disclosure; Page 180-182; 246pp; English.

XX The present invention relates to a novel method of generating a library
CC of fusion nucleic acids. The method involves providing a computationally-
CC derived library of candidate protein sequences and creating a library of
CC expression vectors containing a fusion nucleic acid having a sequence
CC encoding a nucleic acid modification (NAM) enzyme and a sequence encoding
CC a candidate protein sequence from the library and an enzyme attachment
CC sequence (EAS) that is recognised by the NAM enzyme. The invention also
CC relates to the use of a variety of computation methods including protein
CC design automation (pDA). The method is useful in generating and screening
CC fusion nucleic acids that may be used in treating cancer or infections,
CC in detecting protein-protein interactions, discovery of DNA or nucleic
CC acid binding proteins, protein drug discovery, screening for NAM enzymes
CC with decreased toxicity to the host cells and NAM enzyme/EAS pairs with
CC increased affinity or in pharmacogenetic studies. The invention is also
CC used in gene therapy. The present sequence is Adeno-associated virus 5
CC Rep protein. This sequence is used to illustrate the method of the
CC invention. (Updated on 29-AUG-2003 to standardise OS field)

XX SQ Sequence 610 AA;

Query Match 100.0%; Score 2122; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.6e-188;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALVNLVVEHGTTSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDYL 60
Db 221 MALVNLVVEHGTTSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDYL 280
QY 61 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKTNIA 120
Db 281 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKTNIA 340
QY 121 EAIHTVPFYGCNVNTNENFPNDQVDMLIWEEGKMTNKVVEAKAILGSKVRVDQK 180
Db 341 EAIHTVPFYGCNVNTNENFPNDQVDMLIWEEGKMTNKVVEAKAILGSKVRVDQK 400
QY 181 CKSSVQIDSTPVIWTSNTNMVVDGNSSTTFHQQLPDRMFKFELTKRLPPDFGKITKQ 240
Db 401 CKSSVQIDSTPVIWTSNTNMVVDGNSSTTFHQQLPDRMFKFELTKRLPPDFGKITKQ 460
QY 241 EVKOFFAWAKVQNPVTHFEKVPRELAGTKGAESLKRPLGVDVNTSYKSLKRLSPV 300
Db 461 EVKOFFAWAKVQNPVTHFEKVPRELAGTKGAESLKRPLGVDVNTSYKSLKRLSPV 520
QY 301 PETPRSSDVTVPAPLRLPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGNKGCICHNV 360
Db 521 PETPRSSDVTVPAPLRLPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGNKGCICHNV 580
QY 361 THCQICHGIPPEKENLSDFGDDANKQ 390
Db 581 THCQICHGIPPEKENLSDFGDDANKQ 610

RESULT 8

AAE26940 standard; protein; 610 AA.

XX AAE26940;

XX 13-DEC-2002 (first entry)

XX Adeno associated virus 5 Rep protein.

XX Prokaryotic library; candidate protein; nucleic acid modification; NAM;
KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;
KW enzymology; cosmetic research; toxic; environmental safety assessment;
KW nutrient biology; Rep protein.

XX Adeno associated virus.

XX WO20026653-A2.

XX 29-AUG-2002.

XX 14-DEC-2001; 2001WO-US049058.

XX 14-DEC-2000; 2000US-0256163P.

XX (XENC-) XENCOR INC.

XX Li M, Liu Y;

XX WPI; 2002-667068/71.

XX N-PSDB; AAD44600.

XX New library of prokaryotic pET-24a expression vectors, host cells or
PT nucleic acid/protein conjugates, useful for screening candidate proteins
PT and their nucleic acids or modification enzymes for pharmacogenetic
PT analysis.

XX Disclosure; Fig 21; 127pp; English.

XX The invention relates to methods and compositions for the construction of
CC prokaryotic libraries expressing candidate proteins and the use of these
CC libraries to identify candidate proteins and the nucleic acids encoding
CC them. The invention provides a library of prokaryotic pET-24a vectors
CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a
CC nucleic acid modification (NAM) enzyme or a candidate protein, or a
CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or
CC the candidate protein, and an enzyme attachment sequence (EAS) recognised
CC by the NAM enzyme. The library is used for identifying candidate proteins
CC and nucleic acids encoding these proteins, in screening for NAM enzymes
CC with decreased toxicity for the host cells, or in identifying novel or
CC improved EASs, which may be used for understanding cellular processes or
CC any subsequent therapeutic or toxic activities. The nucleic acid/protein
CC (NAP) conjugates are useful in diagnostic assays and in research
CC including clinical pharmacology, functional genomics, pharmacogenomics,
CC agricultural chemicals, environmental safety assessment, chemical sensor,
CC nutrient biology, cosmetic research or enzymology. These may also be used
CC in in vitro screening techniques and in assays with target molecules. The
CC present sequence is Adeno associated virus 5 Rep protein used in the
CC invention

XX SQ Sequence 610 AA;

Query Match 100.0%; Score 2122; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.6e-188;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVNLVVEHGTTSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDYL 60
Db 221 MALVNLVVEHGTTSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDYL 280
QY 61 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKTNIA 120
Db 281 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKTNIA 340
QY 121 EAIHTVPFYGCNVNTNENFPNDQVDMLIWEEGKMTNKVVEAKAILGSKVRVDQK 180
Db 341 EAIHTVPFYGCNVNTNENFPNDQVDMLIWEEGKMTNKVVEAKAILGSKVRVDQK 400
QY 181 CKSSVQIDSTPVIWTSNTNMVVDGNSSTTFHQQLPDRMFKFELTKRLPPDFGKITKQ 240
Db 401 CKSSVQIDSTPVIWTSNTNMVVDGNSSTTFHQQLPDRMFKFELTKRLPPDFGKITKQ 460
QY 241 EVKOFFAWAKVQNPVTHFEKVPRELAGTKGAESLKRPLGVDVNTSYKSLKRLSPV 300
Db 461 EVKOFFAWAKVQNPVTHFEKVPRELAGTKGAESLKRPLGVDVNTSYKSLKRLSPV 520
QY 301 PETPRSSDVTVPAPLRLPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGNKGCICHNV 360

DB 521 PETPRSSDVTVDPAFLPLNWNRSYDCKDYHAQFDNISNKCDEYLNKRGKNGICHNV 580
 QY 361 THCOICHGIPPEKENLSDGDFDDANKEQ 390
 DB 581 THCOICHGIPPEKENLSDGDFDDANKEQ 610

RESULT 9
 ABU64865
 ID ABU64865 standard; protein; 610 AA.
 AC ABU64865;
 XX
 DT 14-MAY-2003 (first entry)
 XX
 DE Rep protein sequence from adeno-associated virus 5.
 XX
 KW Rep protein; capture probe; expression vector;
 KW nucleic acid protein conjugate; NAP; enzyme attachment sequence; EAS;
 KW biochip; gene expression profiling; mutation detection; Rep68; Rep78;
 KW nonstructural protein; NS1; major coat protein; U94.
 XX
 OS Adeno-associated virus 5.
 XX
 PN US2002172968-A1.
 XX
 PD 21-NOV-2002.
 XX
 PF 19-FEB-2002; 2002US-00080376.
 XX
 PR 22-FEB-2001; 2001US-00792630.
 XX
 PA (LIUH/) LIU H.
 PA (DAHI/) DAHIYAT B I.
 PA (LIMM/) LI M.
 XX
 PI Liu H, Dahiyat BI, Li M;
 XX
 DR WPI; 2003-310986/30.
 DR N-PSDB; ABX96669.
 XX
 PT New composition comprising a substrate consisting of an array of capture
 PT probes hybridized to an expression vector or to a nucleic acid protein
 PT conjugate, useful for diagnostic test, gene expression profiling or
 PT mutation detection.
 XX
 PS Disclosure; Fig 21; 125pp; English.
 XX
 CC The invention relates to a composition comprising a substrate comprising
 CC an array of capture probes hybridized to an expression vector or to a
 CC nucleic acid protein conjugate. The capture probes are hybridized to a
 CC expression vector or to a nucleic acid protein (NAP) conjugate. The
 CC vector comprises: (a) a fusion nucleic acid; (b) a capture sequence; and
 CC (c) an enzyme attachment sequence (EAS). The NAP conjugate comprises: (a)
 CC a fusion polypeptide; and (b) an expression vector. The fusion nucleic
 CC acid comprises a nucleic acid encoding the NAP enzyme or candidate
 CC protein. The fusion polypeptide comprises a Rep enzyme and candidate
 CC protein. The EAS and NAP enzyme are covalently attached. Also included are
 CC detecting the presence of a target analyte in a sample, making biochips,
 CC and making NAP conjugates. The composition is useful for diagnostic
 CC applications, gene expression profiling or mutation detection. The
 CC present sequence represents a viral Rep (or related protein e.g. Rep68,
 CC Rep78, nonstructural protein, NS1, major coat protein or U94 protein) for
 CC use in the composition of the invention
 XX
 SQ Sequence 610 AA;

Query Match 100.0%; Score 2122; DB 6; Length 610;
 Best Local Similarity 100.0%; Pred. No. 8.6e-188;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVNLVHEGHTSEKOWIQENQBSYLSFNSTGNSRSQIKALDNTKIMSLTKSAVDYL 60

DB 221 MALVNLVHEGHTSEKOWIQENQBSYLSFNSTGNSRSQIKALDNTKIMSLTKSAVDYL 280
 QY 61 VGSVPEDISKNRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKTNIA 120
 DB 281 VGSVPEDISKNRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKTNIA 340
 QY 121 EAIATHTVPFYGCVANWTNENPFNDKMLIWEEGKMTNKVBSAKAILGSKVRVDQK 180
 DB 341 EAIATHTVPFYGCVANWTNENPFNDKMLIWEEGKMTNKVBSAKAILGSKVRVDQK 400
 QY 181 CKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQOPLDREMFKELYKLPDFGKITQ 240
 DB 401 CKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQOPLDREMFKELYKLPDFGKITQ 460
 QY 241 EVKDFPAWAKVNOQVPTHEFKVPRELAGTGAEKSLKRPGLDVTNTSYKSLKARLSFV 300
 DB 461 EVKDFPAWAKVNOQVPTHEFKVPRELAGTGAEKSLKRPGLDVTNTSYKSLKARLSFV 520
 QY 301 PETPRSSDVTVDPAFLPLNWNRSYDCKDYHAQFDNISNKCDEYLNKRGKNGICHNV 360
 DB 521 PETPRSSDVTVDPAFLPLNWNRSYDCKDYHAQFDNISNKCDEYLNKRGKNGICHNV 580
 QY 361 THCOICHGIPPEKENLSDGDFDDANKEQ 390
 DB 581 THCOICHGIPPEKENLSDGDFDDANKEQ 610

RESULT 10
 ABU64760

ID ABU64760 standard; protein; 610 AA.

AC ABU64760;

XX
 DT 14-MAY-2003 (first entry)
 XX

DE Adeno associated virus a Rep protein, #2.
 XX

KW Biochip; capture probe; nucleic acid modification enzyme; NAM;
 KW enzyme attachment sequence; EAS; single-nucleotide polymorphism; SNP;
 KW protein-protein interaction.
 XX

OS Adeno associated virus 5.

XX
 PN US2002168640-A1.

XX
 PD 14-NOV-2002.

XX
 PF 22-FEB-2001; 2001US-00792630.

PR 22-FEB-2001; 2001US-00792630.

XX
 PA (LIMM/) LI M.

XX
 PA (DAHI/) DAHIYAT B I.

XX
 PI Li M, Dahiyat BI;

XX
 DR WPI; 2003-298722/29.
 DR N-PSDB; ABX96524.

PT Biochip composition useful for creating protein biochips for detecting
 PT target analyte in a sample, has substrate having array of capture probes
 PT hybridized to nucleic acid/protein conjugate.
 XX

PS Disclosure; Fig 21; 123pp; English.

XX
 CC The invention discloses a biochip composition comprising a substrate
 CC having an array of capture probes, which are hybridized to a nucleic acid
 CC (NA)/protein (NAP) conjugate containing a fusion polypeptide, comprising
 CC a NA modification (NAM) enzyme and a candidate protein, and an expression
 CC vector, comprising a NA encoding NAM enzyme and a candidate protein
 CC fusion, a capture sequence and enzyme attachment sequence (EAS). The
 CC biochip composition is useful for detecting the presence of a target

CC analyte in a sample, by contacting the sample with a biochip comprising
CC the compositions under conditions where target analytes can bind to at
CC least one of the candidate proteins to form an assay complex and
CC detecting the presence of target analyte on the substrate. The target
CC analyte is labelled with a fluorescent label and the method further
CC comprises adding a labelled soluble binding ligand to the assay complex.
CC The biochip compositions are useful for creating protein biochips which
CC are useful in diagnosing (detecting the presence of specific target
CC analytes), screening (looking for target analytes that bind to specific
CC proteins), and single-nucleotide polymorphism (SNP) analysis. The bioassay
CC chips are used in assays to determine protein-protein interactions. The
CC target analyte can be nucleic acid, drug, drug analogues or prodrugs. The
CC biochip compositions allow rapid and facile creation of protein biochips.
CC The sequences presented in ABU64750-ABU64772 are the proteins disclosed
CC in the invention
XX

XX SQ Sequence 610 AA;
Query Match 100.0%; Score 2122; DB 6; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.6e-188;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALVNLVEHGITSEKQWIQENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDYL 60
DB 221 MALVNLVEHGITSEKQWIQENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDYL 280
QY 61 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKNTIA 120
DB 281 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKNTIA 340
QY 121 EAIHTVPFYGCNVNTNENFPNDVCDKMLIWEEGKMTNKVESAKAILGSKVRVDQK 180
DB 341 EAIHTVPFYGCNVNTNENFPNDVCDKMLIWEEGKMTNKVESAKAILGSKVRVDQK 400
QY 181 CKSSVQIDSTPVIIVTNTNMVVDGNSSTTFHQQPLEDRMFKELTKRLPDPFGKITKQ 240
DB 401 CKSSVQIDSTPVIIVTNTNMVVDGNSSTTFHQQPLEDRMFKELTKRLPDPFGKITKQ 460
QY 241 EVKDFFAKVNQVPTVTHEFKVPRELAGTKGAESLKRPLGVDVNTNTSYKSLEKRLRSLFV 300
DB 461 EVKDFFAKVNQVPTVTHEFKVPRELAGTKGAESLKRPLGVDVNTNTSYKSLEKRLRSLFV 520
QY 301 PETPRSSDVTVDPAPLRPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGKNGCICHNV 360
DB 521 PETPRSSDVTVDPAPLRPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGKNGCICHNV 580
QY 361 THQIQICHGIPPEKENLSDFGDFDANKEQ 390
DB 581 THQIQICHGIPPEKENLSDFGDFDANKEQ 610

RESULT 11
ABR43398
ID ABR43398 standard; protein; 610 AA.
XX
AC ABR43398;
XX
DE 21-JUL-2003 (first entry)
XX
XX Adeno-associated virus 5 Rep protein SEQ ID NO:21.
XX
XX Library; nucleic acid conjugate; protein conjugate; fusion protein;
KW nucleic acid modification enzyme; candidate protein; screening;
KW enzyme attachment sequence; detection; target analyte; DNA technology;
KW bioinformatic; identification; Adeno-associated virus.
XX
OS Adeno-associated virus 5.
XX
XX WO2003025154-A2.
XX
XX 27-MAR-2003.
XX
XX 12-MAR-2002; 2002WO-US007466.
XX PF

XX 14-SEP-2001; 2001US-00953351.
XX (XENC-) XENCOR.
XX Doberstein SK, Jin CH, Li M, Liu H, Melander C;
XX WPI; 2003-363143/34.
XX N-PSDB; ACC69245.
XX
XX New libraries of nucleic acid/protein (NAP) conjugates or genetic
XX libraries encoding enzyme fusion proteins, useful in DNA technology and
XX bioinformatics, particularly for identifying genes, proteins or analytes.
XX Disclosure; Fig 21; 113pp; English.

XX The present invention describes a library of nucleic acid/protein (NAP)
XX conjugates each comprising: (a) a fusion polypeptide comprising: (i) a
XX nucleic acid modification (NAM) enzyme; and (ii) a candidate protein; (b)
XX an expression vector having a fusion nucleic acid comprising a nucleic
XX acid encoding: (i) the NAM enzyme; and (ii) the candidate protein, where
XX at least two of the candidate proteins are different; and (c) an enzyme
XX attachment sequence (EAS) which is an RNA sequence. The EAS and the NAM
XX enzyme are covalently attached. Also described: (1) a library of
XX expression vectors; (2) making a library of fusion polypeptides; (3)
XX detecting the presence of a target analyte in a sample; and (4) screening
XX a library of small molecules. The library of NAP conjugates is useful for
XX detecting the presence of a target analyte in a sample, or for screening
XX a library of small molecules. The library is useful in DNA technology and
XX bioinformatics, particularly for identifying nucleic acids, proteins, or
XX their sequences, in their native cellular environment. The present
XX sequence represents the Adeno-associated virus 5 Rep protein, which is
XX given in the exemplification of the present invention

XX SQ Sequence 610 AA;
Query Match 100.0%; Score 2122; DB 6; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.6e-188;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALVNLVEHGITSEKQWIQENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDYL 60
DB 221 MALVNLVEHGITSEKQWIQENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDYL 280
QY 61 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKNTIA 120
DB 281 VGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTKNTIA 340
QY 121 EAIHTVPFYGCNVNTNENFPNDVCDKMLIWEEGKMTNKVESAKAILGSKVRVDQK 180
DB 341 EAIHTVPFYGCNVNTNENFPNDVCDKMLIWEEGKMTNKVESAKAILGSKVRVDQK 400
QY 181 CKSSVQIDSTPVIIVTNTNMVVDGNSSTTFHQQPLEDRMFKELTKRLPDPFGKITKQ 240
DB 401 CKSSVQIDSTPVIIVTNTNMVVDGNSSTTFHQQPLEDRMFKELTKRLPDPFGKITKQ 460
QY 241 EVKDFFAKVNQVPTVTHEFKVPRELAGTKGAESLKRPLGVDVNTNTSYKSLEKRLRSLFV 300
DB 461 EVKDFFAKVNQVPTVTHEFKVPRELAGTKGAESLKRPLGVDVNTNTSYKSLEKRLRSLFV 520
QY 301 PETPRSSDVTVDPAPLRPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGKNGCICHNV 360
DB 521 PETPRSSDVTVDPAPLRPLNWSRYDCKDYHAQFDNISNKCDECEYLNRGKNGCICHNV 580
QY 361 THQIQICHGIPPEKENLSDFGDFDANKEQ 390
DB 581 THQIQICHGIPPEKENLSDFGDFDANKEQ 610

RESULT 12
ADI40280
ID ADI40280 standard; protein; 610 AA.
XX

AC ADI40280;
 XX 22-APR-2004 (first entry)
 XX Adeno-associated virus 5 Rep protein SEQ ID NO:27.
 XX protein-binding profile; toxicology; AAV5; Rep.
 XX Adeno-associated virus 5.
 XX WO200268698-A2.
 XX 06-SEP-2002.
 XX 22-FEB-2002; 2002WO-US008023.
 XX 22-FEB-2001; 2001US-0270781P.
 XX (XENC-) XENCOR INC.
 XX Dahiyat B, Li M;
 XX WPI: 2003-040516/03.
 XX N-PSDB; ADI40281.
 XX Creating protein-binding profile of test compound, to create
 XX toxicological profile indicative of compound's toxicity in vivo, involves
 XX screening expression library to identify proteins that interact with
 XX compound.
 XX Disclosure; SEQ ID NO 27; 109pp; English.

CC The invention relates to a novel method for creating a protein-binding
 CC profile of test compound (T) comprising contacting (T) with library of
 CC nucleic acid (NA)/protein conjugates (I), where each (I) has a fusion
 CC protein with NA modifying enzyme (II), and candidate protein (III), and
 CC an expression vector with fusion NA comprising coding sequences for (II)
 CC and (III), and enzyme attachment sequence (EAS). The method of the
 CC invention is useful for carrying out species-specific toxicology tests,
 CC differential organ interaction tests, developmental stage-specific
 CC toxicity tests, and individualised toxicity tests. The method allows for
 CC the creation of a protein-binding profile of the test compound e.g., any
 CC synthetic or natural compound including an organic or inorganic compound,
 CC a peptide or nucleic acid, a metabolite, a drug derivative, or a chemical
 CC entity. The compound can be drug, drug candidate, or an ingredient in
 CC human consumable (e.g., food, textile, cosmetics, flavours, fragrances,
 CC emulsifiers, surfactants, and detergents); compounds that in come in
 CC contact with humans and other animals (e.g., pets and farm stock) such
 CC as pesticides, fertilizers, feed additives, antibiotics, herbicides,
 CC fungicides, polymer additives, and environmental proteins. These profiles
 CC can then be used to evaluate and predict toxicity and other biological
 CC activities of the test compounds. The toxicity profiling information can
 CC be used to predict toxicity of a compound in a different species, e.g.,
 CC to extrapolate the toxicity effects of a compound from one species to
 CC another. The information can also be used to predict the toxicity of a
 CC compound in individuals of the same species that are under different
 CC physiological conditions (e.g., age, sex, and/or disease states, e.g., to
 CC identify individuals particularly susceptible to the toxicity.
 CC Preferably, the method is used to assess the safety of a candidate drug
 CC prior to clinical trial, and to improve clinical trials by allowing
 CC determination of toxic or unanticipated responses in humans early in a
 CC clinical trial to avert tissue toxicity. The methods also helps evaluate
 CC and predict toxicity of chemicals in environmental or occupational
 CC settings such as in manufacturing and agriculture. The present sequence
 CC is used in the exemplification of the invention. Note: The sequences
 CC given in the figures of the specification (ADI40260-ADI40307) are wrongly
 CC labelled in the figure legends. Protein sequences are stated as being
 CC nucleic acids, and vice versa.

XX Sequence 610 AA;
 XX Query Match 100.0%; Score 2122; DB 7; Length 610;
 XX Best Local Similarity 100.0%; Pred. No. 8.6e-188;

	Matches	390;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Cy	1	MALVNLVHGITSEKOWIOENQESYLSNSTGNSRQIKAAALDNATKIMSLTSAVDYL	60							
Db	221	MALVNLVHGITSEKOWIQENQESYLSNSTGNSRQIKAAALDNATKIMSLTSAVDYL	280							
Cy	61	VGSSVPEDISKNRITQIFEMNGYDPAYAGSILYGCQSFNKRNTVMWLYGPAATGKTNIA	120							
Db	281	VGSSVPEDISKNRITQIFEMNGYDPAYAGSILYGCQSFNKRNTVMWLYGPAATGKTNIA	340							
Cy	121	EALAHVFFYGCNVNWTNENPFNDVCDKMLIWEHGGTNNKVVESAKAILGGSKVRVDQK	180							
Db	341	EALAHVFFYGCNVNWTNENPFNDVCDKMLIWEHGGTNNKVVESAKAILGGSKVRVDQK	400							
Cy	181	CKSSVQIDSTPVTNTNMCVVVDGNSSTTFHQOPLDREMFELTKRLPPDGGKITKQ	240							
Db	401	CKSSVQIDSTPVTNTNMCVVVDGNSSTTFHQOPLDREMFELTKRLPPDGGKITKQ	460							
Cy	241	EVKDFFAWAKVQVPTHEFKVPRELAGTKGAEKSLKRLPLGDTVNTSYKSLKRLARLSFV	300							
Db	461	EVKDFFAWAKVQVPTHEFKVPRELAGTKGAEKSLKRLPLGDTVNTSYKSLKRLARLSFV	520							
Cy	301	PETPRSSDVTVDPAFLPLNWSRYDCKDYHAQFDNISNKDCCEVYLNKNGKGCICHNV	360							
Db	521	PETPRSSDVTVDPAFLPLNWSRYDCKDYHAQFDNISNKDCCEVYLNKNGKGCICHNV	580							
Cy	361	THQOICHGIPPEKENLSDFGDFDDANKEQ	390							
Db	581	THQOICHGIPPEKENLSDFGDFDDANKEQ	610							

RESULT 13
 ADH58895
 ID ADH58895 standard; protein; 610 AA.
 AC ADH58895;
 XX 25-MAR-2004 (first entry)
 XX Predetermined property method related adeno-associated virus 5 protein.
 XX predetermined property; population; high throughput; evolution;
 XX functional; cis-acting element; adeno-associated virus.
 XX Adeno-associated virus 5.
 XX US2003224404-A1.
 XX 04-DEC-2003.
 XX 24-FEB-2003; 2003US-00375192.
 XX 25-FEB-2002; 2002US-0360085P.
 XX (VEGA/) VEGA M.
 XX (DRIT/) DRITTANTI L.
 XX Vega M, Drittanti L;
 XX WPI: 2004-033955/03.
 XX Producing molecule such as promoters, silencers, having preselected
 XX property, by producing target nucleic acids having target modified
 XX sequence, expressing and screening for altered function in host cells.
 XX Disclosure; SEQ ID NO 7; 53pp; English.

CC The invention relates to a novel method for producing a molecule having a
 CC predetermined property, by producing population of sets of target
 CC functional nucleic acids having target modified functional sequence,
 CC introducing each set of the molecules into host cells and expressing a
 CC protein and screening the sets of encoded proteins to identify the
 CC molecules whose activity is altered, where each such target functional

CC nucleic acid is designated a hit. The method is a high throughput
CC directed evolution of functional nucleic acid molecules, such as cis-
CC acting elements, particularly those that act in complex biological
CC settings. This sequence represents the protein of an adeno-associated
CC virus used in the novel method of the invention.
XX
PS Sequence 610 AA;
XX
Query Match 100.0%; Score 2122; DB 8; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.6e-188;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALVNLVHGHTSEKQWIOENQESYLSFNSTGNSRSQIKALDNTKIMSLTKSAVDYL 60
DB 221 MALVNLVHGHTSEKQWIOENQESYLSFNSTGNSRSQIKALDNTKIMSLTKSAVDYL 280
QY 61 VGSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTGKTNTA 120
DB 281 VGSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTGKTNTA 340
QY 121 EAIHTVPFYGCVNWTNENFPFNDVCDKMLIWEEGKMTNKVVSASAKAILGSKVRVDQK 180
DB 341 EAIHTVPFYGCVNWTNENFPFNDVCDKMLIWEEGKMTNKVVSASAKAILGSKVRVDQK 400
QY 181 CKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTFEHQOPLDRMFKELTKRLPPDFGKITQ 240
DB 401 CKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTFEHQOPLDRMFKELTKRLPPDFGKITQ 460
QY 241 EVKDFFAWAKVQVPTHEFKVPRELACTGKAESLKRPLGDVNTSYKSLKRLSFFV 300
DB 461 EVKDFFAWAKVQVPTHEFKVPRELACTGKAESLKRPLGDVNTSYKSLKRLSFFV 520
QY 301 PETPRSSDVTVDPAPLRPLNWSRYDCKDYHAQFDNITSNKDCCEYLNRGKNGCICHNV 360
DB 521 PETPRSSDVTVDPAPLRPLNWSRYDCKDYHAQFDNITSNKDCCEYLNRGKNGCICHNV 580
QY 361 THQICIGHPWEKENLSDFGDFDANKQ 390
DB 581 THQICIGHPWEKENLSDFGDFDANKQ 610
RESULT 14
AAV58163
ID AAV58163 standard; protein; 330 AA.
XX
AC AAV58163;
XX
DT 07-MAR-2000 (first entry)
XX
DE Adeno associated virus AAV5 Rep40 protein.
XX
KW Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR;
KW promoter; Rep protein; capsid protein; regulation; transduction;
KW replication; chromosomal integration; tissue tropism; cellular receptor;
KW gene therapy; neutralising antibody; erythroid progenitor cell;
KW transduction; cancer; genetic disease; Rep40.
XX
OS Adeno-associated virus 5.
XX
PN WC9961601-A2.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US011958.
XX
PR 28-MAY-1998; 98US-0087029P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Chiorini JA, Kotin RM;
XX
PI WPI; 2000-062707/05.
XX
DR N-PSDB; AAZ49215.
DR

XX
PT Adeno-associated virus 5 based vectors and particles, useful for gene
XX therapy.
PS Claim 16; Page 86; 91pp; English.
XX
XX This sequence represents the Rep40 protein of adeno associated virus type
CC 5 (AAV5). The invention relates to vectors comprising a pair of AAVS
CC inverted terminal repeats (ITRs) with a promoter between the ITRs. The
CC vector may comprise the viral genome, or subregions thereof, including
CC sequences encoding Rep proteins and capsid proteins, and is encapsidated
CC in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613),
CC Rep52 (AAV58168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in
CC regulation of replication and transduction, in addition to the
CC production of progeny genomes. Rep68 and Rep78 are also associated with
CC the stable integration of the viral genome into human chromosomes. The
CC three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3
CC (AAV58162) assemble to form an icosahedral capsid, and differ from each
CC other by the use of alternative splicing and an unusual translation
CC initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2
CC capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are
CC likely to utilise distinct cellular receptors and are serologically
CC distinct. In a gene therapy application, therefore, AAV5 would allow for
CC transduction of a patient who already possess neutralising antibodies
CC either as a result of natural immunological defence or from prior
CC exposure to AAV2 vectors. The vectors may be useful for transducing
CC erythroid progenitor cells or cells lacking heparin sulphate
CC proteoglycans, which is very inefficient with AAV2-based vectors. The
CC vectors may also be useful for transducing cells with a nucleic acid of
CC interest in order to produce cell lines that could be used to screen for
CC agents that interact with the gene product of the nucleic acid of
CC interest. In addition to transduction of other cell types, transduction
CC of erythroid cells would be useful or the treatment of cancer and genetic
CC diseases which can be corrected by bone marrow transplants using matched
CC donors
XX
SQ Sequence 330 AA;
Query Match 81.1%; Score 1721; DB 3; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.4e-151;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALVNLVHGHTSEKQWIOENQESYLSFNSTGNSRSQIKALDNTKIMSLTKSAVDYL 60
DB 1 MALVNLVHGHTSEKQWIOENQESYLSFNSTGNSRSQIKALDNTKIMSLTKSAVDYL 60
QY 61 VGSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTGKTNTA 120
DB 61 VGSVPEDISKRIWQIFEMNGYDPAYAGSILYGCQSFNKRNTVWLYGPATTGKTNTA 120
QY 121 EAIHTVPFYGCVNWTNENFPFNDVCDKMLIWEEGKMTNKVVSASAKAILGSKVRVDQK 180
DB 121 EAIHTVPFYGCVNWTNENFPFNDVCDKMLIWEEGKMTNKVVSASAKAILGSKVRVDQK 180
QY 181 CKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTFEHQOPLDRMFKELTKRLPPDFGKITQ 240
DB 181 CKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTFEHQOPLDRMFKELTKRLPPDFGKITQ 240
QY 241 EVKDFFAWAKVQVPTHEFKVPRELACTGKAESLKRPLGDVNTSYKSLKRLSFFV 300
DB 241 EVKDFFAWAKVQVPTHEFKVPRELACTGKAESLKRPLGDVNTSYKSLKRLSFFV 300
QY 301 PETPRSSDVTVDPAPLRPLNWSR 324
DB 301 PETPRSSDVTVDPAPLRPLNWSR 324
RESULT 15
AAU11408
ID AAU11408 standard; protein; 330 AA.
XX
AC AAU11408;
XX

26-FEB-2002 (first entry)
Adeno-associated virus 5 (AAV5), Rep40 protein.
Adeno-associated virus 5; AAV5; Rep40; neuroprotective; cyfostatic; gene therapy; Parkinson's disease; Alzheimer's disease; demyelination disease; metabolic disorder; musculoskeletal disease; cardiovascular disease; cancer; autoimmune disorder; genetic disease; cystic fibrosis; pseudohypoadosteronism; immotile cilia syndrome; bronchitis; pneumonia; emphysema; pulmonary oedema; central nervous system; replication; transcription.
Adeno-associated virus 5.
WO200170276-A2.
27-SEP-2001.
22-MAR-2001; 2001WO-US009123.
22-MAR-2000; 2000US-00533427.
(IOWA) UNIV IOWA RES FOUND.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Chiorini JA, Kotin RM, Davidson B, Zabner J;
WPI; 2002-055104/07.
N-PSDB; AAS17712.
Delivering nucleic acid into cell for treating Parkinson's disease, by administering to cell an adeno-associated virus 5 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats.
Disclosure; Page 125; 130pp; English.
The invention describes a novel method of delivering a nucleic acid into a cell in a subject, comprising administering to the cell an adeno-associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus which relies on a helper virus for replication, in the absence of which the AAV5 genome is integrated into a host chromosome in a locus specific manner. The method provides a way to deliver a nucleic acid to a specific regions, tissues and cell types of the central nervous system comprising inserting the nucleic acid between a pair of AAV inverted terminal repeats or delivering an AAV5 particle containing a vector comprising the nucleic acid. The method is useful for treating brain disorders such as demyelination disease, Alzheimer's disease and Parkinson's disease, and metabolic disorders such as musculoskeletal diseases, cardiovascular disease, cancer and autoimmune disorders, for treating genetic diseases such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoadosteronism, immotile cilia syndrome, and for treating bronchitis, pneumonia, emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is useful for delivering gene that may have a systematic effect like anti-hypertension drugs, insulin, coagulation factors, antibiotics, growth factors and hormones. This is the amino acid sequence of the adeno-associated virus 5 (AAV5) Rep40 protein, one of 4 Rep proteins that regulate replication and transcription of the AAV5 genome, described in the method of the invention
Sequence 330 AA;
Query Match 81.1%; Score 1721; DB 5; Length 330;
Best Local Similarity 100.0%; Pred.No. 6.4e-151;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALVNWLVHEHGTSEKQWIQENQSYLSFNGTNSRSQIKALDNTKIMSLTKSAVDYL 60
DB 1 MALVNWLVHEHGTSEKQWIQENQSYLSFNGTNSRSQIKALDNTKIMSLTKSAVDYL 60
QY 61 VGSSVPEDISKNRWQIFEMNGYDPAYAGSILYGCQSFNKRNTWLYGPATTKTNIA 120
DB 61 VGSSVPEDISKNRWQIFEMNGYDPAYAGSILYGCQSFNKRNTWLYGPATTKTNIA 120

QY 121 EALAHTEFFYGCYVNWNTNENFFPNDVCKMLIWWBEGKMTNKVVSAGKAILGSGKVRVDQK 180
DB 121 EALAHTEFFYGCYVNWNTNENFFPNDVCKMLIWWBEGKMTNKVVSAGKAILGSGKVRVDQK 180
QY 181 CKSSVQIDSTPVIYVTSNTNMVVDGNSSTTPEHQOPLBDRMFKFELTKRLPPDFGKITQK 240
DB 181 CKSSVQIDSTPVIYVTSNTNMVVDGNSSTTPEHQOPLBDRMFKFELTKRLPPDFGKITQK 240
QY 241 EVKDDFFAWAKVNOVPVTHEPKVPRELAGTKGAEKSLKRLPLGDTVNTSVKSLKRLSFKV 300
DB 241 EVKDDFFAWAKVNOVPVTHEPKVPRELAGTKGAEKSLKRLPLGDTVNTSVKSLKRLSFKV 300
QY 301 PETPRSSDVTVDPAPLRPLNWSR 324
DB 301 PETPRSSDVTVDPAPLRPLNWSR 324

Search completed: October 23, 2004, 20:36:55
Job time : 119 secs

This Page Blank (uspto)